RIVISTA INTERNAZIONALE DI FILOSOFIA E PSICOLOGIA ISSN 2039-4667, E-ISSN 2239-2629
DOI: 10.4453/rifp.2019.0002 Vol. 10 (2019), n. 1, pp. 16-32

RICERCHE

Emotions, Experiments and the Moral Brain. The Failure of Moral Cognition Arguments Against Moral Sentimentalism

Lasse T. Bergmann[(α)]

Ricevuto: 30 novembre 2018; accettato: 29 marzo 2019

Abstract Moral cognition research has in part been taken to be a problem for moral sentimentalists, who claim that emotions are sensitive to moral information. In particular, Joshua Greene can be understood to provide an argument against moral sentimentalism on the basis of neuropsychological evidence. In his argument he claims that emotions are an unreliable source of moral insight. However, the argument boils down to circular claims: (1) Rationalistic factors are assumed to be the only morally relevant factors; (2) Emotions are not sensitive to these factors; (3) Thus, Moral Sentimentalism is false, because only rationalistic factors are justified. While this circularity makes so-called sourcing-arguments fallacious if applied against moral sentimentalism, moral cognition research has much to contribute. Indeed, moral cognition research will be instrumental for clarifying the sentimentalist position, shedding light on the mental mechanics underlying emotional moral processing. After all, evidence from moral cognition points to substantial involvement of affective processes in human moral cognizing and their embodied nature; thus, challenging long held beliefs about morality.
KEYWORDS: Moral Cognition; Moral Sentimentalism; Emotions; Embodied Cognition; Moral Brain

Riassunto *Emozioni, Esperimenti e il cervello morale. L'errore degli argomenti basati sulla cognizione morale contro il sentimentalismo morale* – Si è spesso ritenuto che la ricerca nell'ambito della cognizione morale costituisse, almeno in parte, un problema per il sentimentalismo morale, il quale sostiene che le emozioni sono sensibili all'informazione morale. In particolare, si può pensare che Joshua Greene abbia portato un argomento contro il sentimentalismo morale basato su evidenza neuropsicologica. Secondo il suo argomento le emozioni non costituiscono una fonte affidabile di comprensione morale. E tuttavia questo argomento fa leva su una circolarità: (1) gli unici fattori qualificati come moralmente rilevanti sono quelli razionali; (2) le emozioni non sono sensibili a questi fattori; (3) pertanto, considerato che solo i fattori razionali sono giustificati, il sentimentalismo morale è falso. Tale circolarità rende fallaci i cosiddetti argomenti sorgente, laddove questi siano applicati al sentimentalismo morale. Al contempo, la ricerca sulla cognizione morale ha molto da dire su questo argomento perché può contribuire a chiarire la posizione sentimentalista, gettando luce sulla meccanica mentale sottostante i processi morali che fanno leva su emozioni. Anzi questa prospettiva di ricerca evidenzia che i processi di carattere affettivo sono coinvolti in maniera sostanziale nella cognizione morale umana e hanno una natura incarnata; in questo modo essa mette in discussione convinzioni di vecchia data sulla morale.
PAROLE CHIAVE: Cognizione morale; Sentimentalismo morale; Emozioni; Cognizione incarnata; Cervello morale

[(α)]Institut für Kognitionswissenschaft, Universität Osnabrück, Wachsbleiche 27 - 49090 Osnabrück (D)

E-mail: lasse.bergmann@uos.de (✉)

IN THE LAST DECADE, MORAL cognition research has made some inroads into mainstream philosophy. This ambitious interdisciplinary project aims to understand the cognitive processes involved in human cognizing in morally challenging situations. Understanding these mental processes has wide ranging philosophical implications. Philosophers have speculated for a long time about the workings of moral judgments, beliefs, emotions, motivation, etc. with profound implications for their philosophical work. Theorizing about what it means to live a virtuous life, to be a good person, and to make the right decisions can be enriched by scientifically accurate empirical accounts of the mental mechanics speculated about.

At times, however, moral cognition researchers have even more ambitious goals. Not satisfied by merely settling empirical claims, they attempt to settle normative claims. These claims have traditionally belonged solely to the domain of philosophical inquiry – but at the moment, in the advent of experimental philosophy, attempts are being made to settle normative disputes on the basis of experimental insight. One of the disputes picked up by the field focuses on the role of emotions in moral cognizing: the long-standing disagreement between rationalists and moral sentimentalists on whether emotions pick up on morally relevant facts. This debate goes back to at least Kant and Hume, who disagreed about whether sentiment or reason is the dominant force of the moral mind.[1] More recently some work in moral cognition has been taken to be a problem for moral sentimentalists, e.g., Kauppinen stressed that: «[...] Greene makes a plausible case that some emotional processes are sensitive to morally irrelevant factors [...]».[2]

Joshua Greene should be acknowledged as the most substantial contributor to the field of moral cognition.[3] Certainly, his work has implications for moral sentimentalists, though he is mainly concerned with the entrenched debate between deontology and utilitarianism. He argues that intuitions supporting deontology are merely emotional responses, which do not provide valid epistemic access to moral facts. He constructs this argument by experimentally linking intuitions supporting deontology to emotional responses and at the same time assailing the validity of moral sentimentalism. His argument boils down to this: moral emotional responses are prone to biases and the source of deontological intuitions are such biased emotional responses. Thus, he argues, deontological intuitions should not be considered morally valid intuitions, because they ironically are a result of biased emotional responses.

This argument, which draws on a large body of psychological and neuroscientific research, has attracted a lot of attention. It provides a forceful attack on the position of moral sentimentalists, who claim that emotions inform and motivate ethical decisions. Subsequently, many authors have articulated criticism of Greene's and related arguments by trying to disprove specific argumentative steps.[4] However, it is my contention that Greene provides a fallacious argument to begin with. As such, the correctness of single premises should be of secondary interest, as the argument presented by Greene is circular. The argument assumes a principle about what is the normatively correct outcome of a decision problem, on the basis of which factors should be considered. Through this principle normativity is imbued into cognitive processes, which, in turn, is used to derive normative conclusions. However, the conclusions settle normative questions, which tacitly have already been decided by the starting assumption of the normative principle.

Moral cognition research, however, has other – less ambitious – arguments to offer that have implications for moral sentimentalism. These arguments are non-circular, because they do not rely on a normative first principle, but point towards inconsistencies between normative assertions of theories and the empirical stipulations going along with them. Such arguments are defensible and have already had a lasting effect on philo-

sophical discourse, demonstrating the value of philosophical work informed by moral cognition research. Moral sentimentalists should, in the end, be encouraged to embrace moral cognition research. The insights gained are useful to clarify sentimentalist positions and, more often than not, point to substantial involvement of emotional processes in moral processing[5] – one of the core beliefs of every sentimentalist.

Setting the Stage

Moral cognition research is a fairly recent development and the new breed of arguments created are still in a trial phase. The complexity of this interdisciplinary endeavor that draws from cognitive science and ethics often produces complex arguments. It is first of all instructive to distinguish two research projects, though they are often pursued at the same time: (1) Experimental philosophy, a project that tries to settle philosophical debates by constructing them into experimentally testable forms and deriving philosophical conclusions from experimental results; (2) Moral cognitive-(neuro-)psychology, a project investigating the nature of cognitive processes involved in morally challenging tasks, i.e., moral decision-making, moral motivation, and moral perception. Often these projects are not suitably distinguished from one another and the complexity of the arguments obfuscates their content. The moral cognition aspect of such arguments generally takes form along the lines of the following premises:

- Psychological or neuroscientific experimental results suggest that neuropsychological factors X and Y differ in their (moral) validity.
- Intuitions[6] that rely on factors X are valid and those relying on Y are invalid.

These insights can then be put to work to settle debates in moral philosophy (thus dealing with issues relating to experimental philosophy). There are two ways to complete these arguments. Depending on which kind of completion for the argument is formulated, different kinds of arguments can be constructed: *Fortification arguments*, which aim to defend the validity of intuitions:[7]

- Moral Theory M is supported by Intuition I.
- Intuition I relies on factor X.
- Thus, the support for Moral Theory M by Intuition I is valid.

Or alternatively, *debunking arguments*, assailing the validity of an intuition:

- Moral Theory M is supported by Intuition I.
- Intuition I relies on factor Y.
- Thus, the support for Moral Theory M by Intuition I is invalid.

The latter kind are of greater interest for moral sentimentalists,[8] as these kinds of arguments are often employed against sentimentalists to debunk their position,[9] rather than by sentimentalists to fortify their position.

In the following, a specific species of debunking arguments shall be considered, which will be labeled *sourcing arguments*, i.e., arguments that purport the neuropsychological source of intuitions to be the factor relevant to the validity of these intuitions. These arguments assume that different kinds of cognitive processes produce different kinds of moral thought. These kinds of moral thought can be distinguished in terms of validity on the basis of characteristics of the cognitive systems producing this thought. Sourcing arguments can take the form of both debunking and fortification arguments. The concrete exemplary case is taken from Joshua Greene's work. His work generally deals with universalist moral issues and heavily utilizes trolley problems. He has contributed many arguments to the debates in moral cognition, which are sometimes difficult to keep apart due to similarities in their structure. The core of the argument at issue here

has most clearly been stated in Greene's *No Cognitive Miracles Principle*:

> When we are dealing with unfamiliar moral problems, we ought to rely less on automatic settings (automatic emotional responses) and more on manual mode (conscious, controlled reasoning), lest we bank on cognitive miracles.[10]

In a nutshell, this claim distinguishes cognitive processes, one that reliably provides valid moral insight and one which does not. Thus, intuitions, thought, or responses sourced form one process should be trusted and relied upon and those from the other process should not. The argument Greene presents for this claim is complex and has been articulated in many different ways and with different purposes.[11] Thus, in the following, his argument will be reconstructed to reveal its circularity when applied against the position of moral sentimentalism. The first step of reconstruction will be to put it in a general form:

(I) If psychological process T1 (automatic-emotional thinking) is the source of insight, then this insight is unreliable and if psychological process T2 (controlled-rational thinking) is the source of insight, then this insight is reliable.
(II) The source of intuition I1 is T1 and the source of Intuition I2 is T2.
(III Intuition I1 supports moral principle M1 and Intuition I2 supports moral principle M2.
(IV) Thus, the support of intuition I1 for principle M1 can be considered unreliable, while the support of intuition I2 for principle M2 can be considered reliable.

There are two major steps in this argument: (1) Explicating a specific version of the dual-process view inherited from Kahneman, which makes premise (I) tenable;[12] (2) The application of this view to moral disagreements, i.e., empirical work showing that intuitions actually depend on the cognitive process they are sourced from and philosophical work that shows these intuitions support moral theories. While many authors have argued against either of these steps in many ways, the cardinal sin of the argument, i.e., the circularity of this argument, is only revealed by making the normative assumptions more explicit. In this paper, individual premises will not be treated overly critically, even though they might be questionable or worthy of deeper discussion, because its purpose is to investigate the overarching problem of the argument's hidden circularity. To find this hidden element in Greene's argument, the next section will give a brief overview of said argument and how it plays out.

Greene, experimental philosophy and explicit normativity

Greene's sourcing-argument unfolds as follows: Greene shows experimentally that utilitarian intuitions are associated with different neural processes than deontological intuitions. He describes an episode with his early mentor Jonathan Cohen:[13] Greene, since his days as a philosophy PhD student, had harbored the suspicion that moral judgements differ depending on whether they involve personal interactions or not. Cohen quickly connected this to different neuronal pathways, and thus a new fruitful interdisciplinary endeavor was born.

In fMRI experiments, Greene and Cohen isolated the neural processes elicited in tasks that focus on the trolley problem, an influential thought experiment in normative ethics.[14] The thought experiment poses the question whether one would sacrifice one person to save five others. In the version originally introduced by Foot,[15] five railway workers are in the path of a runaway trolley (from which the experiment inherits its name). The question now is whether one would pull a lever to redirect the trolley to another track. The tricky morally relevant detail is that on this alternative track there is a single person

that would be hit by the trolley if one chooses to redirect it. The most significant alternative version of the problem was posed by Thomson.[16] In her variation the five railway workers can be saved from the oncoming trolley by pushing a large man from a footbridge onto the tracks. The body of the large man is heavy enough to stop the trolley, so that the lives of the five workers can be saved in exchange for sacrificing the large man. The interesting feature of these trolley cases is that the different versions elicit different intuitions.

The first case presented (the lever case) usually elicits the intuition that one should pull the lever. This intuition is in line with utilitarian normative demands and contrary to those deontological theories that purport that doing nothing is better than actively killing a person by pulling the lever. The second case (the footbridge case), however, usually elicits the intuition that pushing the large man in front of the trolley is wrong, even though it would save many lives. This is in tension with demands of most utilitarian theories, but this intuition is in line with deontological considerations. This research paradigm mirrors a professional philosopher conducting a thought experiment, that supposedly gives them insight into the validity of a philosophical theory, depending on whether or not the state of affairs in the thought experiment is properly and plausibly accounted for by the philosophical theory. Correspondingly, the logic of experimental philosophy takes experimentally established responses corresponding to claims derived from moral theories as support for this moral theory as such.[17]

Greene's fMRI studies show that intuitions in the lever case are associated with characteristic activation in the dorsolateral prefrontal cortex (dlPFC), an area associated with controlled-rational processing.[18] Based on this insight, Greene says that we should place more importance on this intuition. The intuition in the footbridge case, however, is associated with activation in the ventromedial prefrontal cortex (vmPFC), an area associated with automatic-emotional processing and thus Greene discounts these intuitions. The reason why one cognitive process can be trusted and the other cannot is (ostensibly) provided by experimental work by Kahneman and others. This body of work is the source of the tacitly assumed normativity in this argument, as the following section will reveal.

Kahneman, moral cognition and hidden normativity

Kahneman famously investigated cognitive biases and advanced a theory of two modes of cognition: a fast and a slow mode.[19] This theory and its many descendants are today known as dual-process theories. The idea is simple: by means of many experiments, brain regions have been identified which are more and less involved in producing some "correct" behavior. These experiments revealed that the fast processes were prone to be more involved in undesirable behavior, while the slow processes were usually responsible for the desired outcome. The labels "fast" and "slow" changed over time. The notion of a fast process took on emotional connotations: such as, subjective, heuristic, automatic, unconscious. Slow processes, however, gained an air of infallibility and were given many rationalistic labels: for example, controlled, flexible, objective, deliberate. Kahneman offers a very neutral terminology, calling fast processes "type 1 processes" and slow processes "type 2 processes". His more normative labels stuck, however, and the use of dual-process theories became normatively charged. The claim that type 2 processing (slow-controlled, i.e., rational) is "good" (normatively correct) and type 1 processing (fast-automatic, i.e., emotional) is "bad" is well rehearsed in dual-process accounts of cognition.[20]

Evans brands this normative connotation fallacious and diagnoses it as an overgeneralization of classic experimental results:

> In the traditional paradigms, researchers presented participants with hard, novel

> problems for which they lacked experience (students of logic being traditionally excluded), and also with cues that prompted type 1 [fast-automatic] processes to compete or conflict with these correct answers. So in these paradigms, it does seem that type 2 [slow-controlled] processing is at least necessary to solve the problems, and that type 1 [fast-automatic] processes are often responsible for cognitive biases.[21]

This means that the standard way to conduct these experiments leads to experimental setups in which only one sort of reasoning (controlled) can lead to correct responses, and other kinds of reasoning can only produce errors. Evans continues that in more recent experimental work, in which a broader range of cognitive tasks is considered, both types of processing seem to be equally responsible for cognitive biases.[22]

Evans shows that Kahneman's insights are domain specific (at best) and if a more comprehensive domain of cognitive tasks is considered neither type 1 nor type 2 processing provides some sort of normative advantage. Thus, premise (I) is empirically inadequate and needs to be revised. Crucially, Greene's argument assumes that processes T1 and T2 are *generally* reliable or unreliable sources of insight. However, these processes appear to have limited flexibility and can thus only be considered reliable in a specific domain. These cognitive processes then turn out to be unreliable in different domains. This means that the experimental results are overgeneralized. I do not take this point to be a fatal flaw in the argument that cannot be remedied by adapting the argumentative structure. The argument from above, specifically premise (I), can be reformulated to account for this:

(I*) In domain D, if psychological process T1 (type 1 automatic-emotional thinking) is the source of insight, then this insight is unreliable and if psychological process T2 (controlled-rational thinking) is the source of insight, then this insight is reliable.

(D) Intuition I is only elicited in domain D.

In this formulation of the argument, it depends on the truth of premise (D) whether the argument is overgeneralizing or not. There, however, is still another problem with premise (I*). Evans points to an implicit correctness criterion used to conduct the experiments that establish the reliability of the different types of processing: «The problem is that a normative system is an externally imposed, philosophical criterion that can have no direct role in the psychological definition of a type 2 process».[23]

Since this philosophical criterion has never been explicitly stated and has been implicitly used by a multitude of researchers, its footprint in the research is probably fuzzy, but nevertheless it has a distinctly rationalistic connotation. Experiments often take the form of logic puzzles «like assuming some dubious proposition to be true and deciding whether a conclusion necessarily follows from them». [24] Kahneman's examples to introduce the flaws of type 1 processes are all drawn from mathematics or logic.[25] These examples present problems that are supposed to reveal the flaws of fast processing:

- A bat and ball cost $1.10. The bat costs one dollar more than the ball. How much does the ball cost?
- All roses are flowers. Some flowers fade quickly. Therefore, some roses fade quickly. Is this a valid inference?

The quick response of many people, especially those untrained in these kinds of puzzles, will be wrong: Their first response may be that the ball costs 10 cents or that the conclusion of the second example can be inferred from the premises. In such logic puzzle experiments the general truth of the axioms of classical logic is assumed and given the task this assumption is almost too trivial to mention. But it is a normative criterion and the

triviality of applying it to these examples should not be taken to mean that it is a trivially true normative criterion. If moral philosophy tells us anything, it is that there are no trivially true normative criteria. Consider the following experiment, where subjects get to know the fictional character Linda:

> Linda is thirty-one years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.[26]

Then they are asked to evaluate the likelihood of a number of statements, two of which are:

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

Most subjects rate the second statement as more likely, which is nonsense in terms of probability theory (conjunctions cannot be more likely than their conjuncts). This failure to exhibit the right behavior, i.e., to give the right answer, can be connected to fast processing. But it cannot be said that the fast processing picked up on something meaningless here. Given the description of Linda, the second statement makes more sense, i.e., presents a more harmonious picture of a person. While it is clear that the fast response *did not read the instructions carefully*, it picked up on something.

What is at issue here is first and foremost how normativity is imbued into a cognitive process. Can there be a better or worse process?

It seems reasonable to accept that there are processes that are better or worse at doing certain jobs, e.g., controlled processes are more likely to be involved in correctly solving logic puzzles in untrained subjects. But which job would a process need to be good at so that it can be considered a morally good process? Furthermore, is there no potentially morally relevant job that fast processing is suited for? There are certainly philosophical issues in question here that cannot be settled on merely empirical grounds. Those philosophical issues are precisely the normative issues that the sourcing argument tries to settle. Some normative theories will postulate that the relevant job is to follow the demands of the principle of utility, other normative theories posit sensitivity to different features in the world to be the relevant job. Whether a process is *good or bad* is settled based on whether it does the relevant job. Whether the job is the relevant one depends on its sensitivity to morally relevant factors, which in turn depends on which normative theory is endorsed – an issue that needs to be settled on philosophical grounds. Thus, to impose a normative criterion on cognitive processes is primarily a philosophical endeavor. To provide examples (i.e., experimental evidence) which show that a certain process is good at a certain job says nothing about whether this is the morally relevant job. This means that the correctness of any premise establishing the reliability of a cognitive process involved in moral cognition depends on which ethical theory is favored, the precise question sourcing arguments purport to settle.

Certainly, it is convincing that automatic processes of untrained subjects are prone to biases in solving logical puzzles, as, apparently, they are prone to pick up on other features of the world. But it is an entirely different question whether this a morally relevant job. Moral rationalists liken the truths of morality to those of mathematics and logic: according to them, they are a priori truths. Thus, they consider the cognitive task of "detecting" a priori truths in the world morally relevant. It is such an a priori truth, that it is not more likely for Linda to be a bank teller and a feminist than it is for her to be a bank teller. So, the rationalist may believe that type 2 processes are "good," because they do the job of applying some a priori knowledge to the world. However, sentimentalists do not share such a commitment. After all, if emotions are

the source of moral knowledge, this knowledge is presumably a posteriori. The unintentional strategy that has been tacitly employed in the sourcing argument is to assume, through the use of cherrypicked situations and tasks, a normative criterion. In these cases, the correctness of mathematics or logic seems self-evident, thus a principle favoring behaviors indicating the mathematically or logically correct solution is trivially applicable in these cases. This hides the fact that all that has been shown is that these processes are better at doing a specific job and are only *better* processes in virtue of the assumption of principles proclaiming this to be the relevant job – a principle that is rationalistic in nature.

The Circularity of Sourcing Arguments

As has been illustrated in the paragraphs above, the reliability of cognitive processes has been established on the basis of a rationalistic principle: a principle C saying, that, for all situations in domain D, if a desired response (in these cases that responses in accordance with mathematics and logic) is achieved by P, then P is a reliable source of insight. So premise (I*) can be restated as follows:

(I**) As established through behavioral principle C, in domain D, if psychological process T1 (automatic-emotional thinking) is the source of insight, then this insight is unreliable and if psychological process T2 (controlled-rational thinking) is the source of insight, then this insight is reliable.

While the last section showed the structure of sourcing arguments to be circular, it is worth discussing how such sourcing arguments can be employed against various positions. Based on the clarifications developed in the last sections, Greene's sourcing argument can be stated as follows:

(i) As established by principle C, in domain D, if psychological process T1 (type 1 automatic-emotional thinking) is the source of insight, then this insight is unreliable and if psychological process T2 (type 2 controlled-rational thinking) is the source of insight, then this insight is reliable.
(ii) Intuition I1 supporting deontology is sourced from T1.
(iii) Intuition I2 supporting utilitarianism is sourced from T2.
(iv) The intuitions in question are elicited in domain D.
(v) Thus, from (i), (ii) and (iv) it can be concluded that intuition I2 supporting utilitarianism is reliable and should be trusted. From (i), (iii) and (iv) it can be concluded that intuition I1 supporting deontology is unreliable and can be discounted.

The objection this paper is committed to is that sourcing arguments boil down to circular claims. The sourcing argument is employed to settle normative debates; that one normative principle is preferable over another. However, premise (i) already presupposes such a conclusion, since a specific kind of normative principle is employed. As has been argued above, in the process of imposing normative characteristics on a cognitive process it is presupposed which facts or information are morally relevant (see Section 4). To make such a determination, however, is precisely what moral theories usually offer. To conclude that one moral theory is superior to another is also to accept the assertion the theory makes about which information is morally relevant.

Consider the following polemic formulation of the argument, which is intended to highlight the critical flaws of the argument:

1. Morally relevant information is only contained in the logical or mathematical features of the world.[27]
2. Emotions (or more generally affective processes) are sensitive to non-logical and non-mathematical features of the world.

3. Moral intuitions relying on emotional processes lack moral validity, as the processes that produce them are not sensitive to morally relevant information.
4. Normative theory M is based on intuitions relying on emotional processes.
5. Thus, normative theory M, holding non-logical or non-mathematical features of the world to be morally relevant, lacks morally valid support and should therefore be rejected.

Through premise (1) a normative commitment in favor or moral rationalism is made. Thus, premise (1) already presupposes the falseness of theory M, qua this theory holding different factors to be morally relevant than those assumed to be the correct ones. Thus, such an argument is circular if employed against theories holding other features to be relevant than those assumed in premise (1). One significant aspect of the debate that this sourcing argument purports to settle is which information is morally relevant. Utilitarians tend to view moral problems as calculable and it is hardly surprising that cognitive processes sensitive to mathematically relevant information will produce moral thought that is in line with utilitarian considerations.

Other normative theories, however, attribute no significance to how many people are affected by a decision. Deontology, for example, assigns critical importance to whether persons are used as mere means, irrespective of the amount of harm potentially avoided by using them. Such considerations presumably stem from other cognitive processes. A deontologist should thus not agree that the mathematical features of the world are the bearers of morally relevant information. It is somewhat ironic should a deontologist not agree to a rationalistic criterion, as the father of deontology, Immanuel Kant, is considered a rationalist par excellence. But this may simply be the wrong flavor of rationalism.

The more intriguing possibility, that deontology may not be as rational a theory as it purports itself to be, will be developed in the next section, in the context of another species of arguments in moral cognition.

However, whether a *sourcing argument* is spelled out to assail moral sentimentalism, or defend utilitarianism or rationalism is of no consequence, all these claims are in the end circular. A principle built on the truth of mathematics or logic is essentially rationalistic. Hence, it will likely exclude conclusions favoring moral sentimentalism. In the previous passages, the formulation of Greene's claim was more explicitly directed at his main interest, the debate between deontology and utilitarianism. As his argument is taken to be a problem for moral sentimentalists I will be spelled out as such once:

(i) As established by a rationalistic behavioral principle, in domain D, if psychological process T1 (type 1 automatic-emotional thinking) is the source of insight, then this insight is unreliable.
(ii) Intuition I, picking up on factors deemed relevant by moral sentimentalists and thus supporting moral sentimentalism, is sourced from T1.
(iii) The intuitions in question are elicited in domain D.
(iv) Thus, it can be concluded that intuitions supporting moral sentimentalism are unreliable and can be discounted according to a rationalistic principle.

As far as we can treat rationalism as an antithesis to moral sentimentalism, this argument is circular. Rationalism is assumed to be true in premise (i) and thus the conclusion that the antithesis is false, is a circular claim. However, there are a number of positions that may present a partial synthesis of rationalism and sentimentalism, as such this formulation of the argument may not capture them adequately. However, what is meant here with rationalism is almost certainly that the truths of morality are akin or similar to the truths of mathematics and logic. A claim which, I believe, is almost univer-

sally a problem for sentimentalists – I take it as somewhat obvious that emotions cannot be trusted to produce mathematically or logically correct responses. If a "good" brain process is one that produces "right" responses according to the principles of mathematics and logic, then moral sentimentalism is assumed to be false to begin with. And as this appears to be a premise of Greene's argument, it cannot be applied to conclude the falseness of moral sentimentalism.

Debunking without sourcing

The considerations advanced here should not be misunderstood as a criticism of moral cognition research or experimental philosophy in general, which are in many instances very promising research endeavors. They do, however, fail to provide a non-circular argument against moral sentimentalism proper. Nevertheless, specific sentimentalist positions may fall prey to experimental insights.

One of the problems with the work in moral cognition is that arguments bleed into each other. In contrast to many sourcing arguments, debunking arguments can be constructed in a non-circular way; though elements of the sourcing argument are sometimes implicitly assumed to be part of these arguments, without actually including them. Therefore, ostensibly implying the falseness of moral sentimentalism, but not doing the required argumentative work. The most common way to clarify (one of) Greene's argument(s) is to cash out the experimental data he provides in the following form:[28]

1. Deontology holds factor X to be morally relevant.
2. Moral agents' preferential support for actions in accordance with deontology are caused by the agents picking up on factor X.
3. Experimental data shows that factor X is actually irrelevant for the intuition, which is elicited by factor Y.
4. Factor Y is not held to be morally relevant by Deontology.
5. Thus, moral agents' preferential support of deontology has been debunked.

This is both a valid argument and is not vulnerable to the circularity objection discussed before. This argument does not rely on externally imposing a normative criterion on a cognitive process, it relies on a moral theory stipulating certain factors to be relevant and then experimentally revealing that these factors do not play the role they are supposed to be playing. This means that the theory makes assertions about moral cognizing which are inconsistent with experimental evidence giving insight into these cognitive processes.

Based on this analysis, I will label this species of debunking arguments *inconsistency arguments*. For example, one could construct an argument against deontology on the basis of Kant's rejection of emotional experience as a morally relevant factor. A deontologist joining Kant in this commitment may rightfully be troubled by the experimental evidence that emotions seem to play a critical role in moral cognition in general or the role of emotions in eliciting characteristically deontological intuitions. Such a rationalistic commitment is inconsistent with the experimental insights. However, this should not be taken to mean that the position of moral sentimentalism is in any kind of trouble. Though it appears to be a common theme that cognitive processes do pick up on factors not endorsed by moral theories and these cognitive processes are often asserted to be emotional processes, it seems somewhat unclear whether this is a problem for moral sentimentalism. First of all there is the conceptual question of whether these processes actually should be considered emotional or moral emotions proper, and second the question of whether the factors they pick up on are truly morally irrelevant.[29]

Inconsistency arguments have a more limited argumentative force than the sourcing argument discussed in the beginning. While these arguments are more defensible and do

a reasonable job of attacking specific moral doctrines, e.g., deontology, they provide little reason to doubt moral sentimentalism. While a rationalist may be troubled by the association of their moral intuitions with emotional processes, this certainly is not a surprise or worry to a sentimentalist.

The best of both worlds (or how some moral cognition arguments may have limited force against moral sentimentalism)

The presumably most impactful arguments emanating from moral cognition research on the field of moral sentimentalism will be a species of inconsistency arguments, which experimentally reveal the nature of emotional moral processing to be contrary to philosophical assumptions. Moral sentimentalism has to provide conceptions of emotions that make their claim that emotions pick up on morally relevant factors plausible. As emotions are some sort of mental process, emotion theorists have to make empirical assertions, which are subject to experimental investigation. One of the substantial insights of Greene's work appears to be that emotional moral responses are bound to sensory-motor programs sensitive to context. Greene and colleagues write: «In a general sense, this suggests a mechanism of moral judgment that is a species of embodied cognition».[30] This conclusion is based Greene's work previously discussed, but draws more specifically on the empirical insights gained therein.

One much discussed factor that influences agents' judgements but is not prima facie endorsed by moral theory (i.e., deontology), can be precisely isolated in different variations of the footbridge dilemma. Deontologists hold that the footbridge dilemma elicits a characteristically deontological response, i.e., not pushing the fat man to save the others, because they would not use a single person to save the other, i.e., using a person as a means to an end. This explanation would also account for agents' willingness to push the lever in the original case, as the person sacrificed there is not used as a means to save the other potential victims. Greene and colleagues find that this is, however, not the relevant factor that people respond to. If, for example, the fat man is dropped onto the track not by pushing, but by a trap-door connected to a mechanism that is triggered by a lever, then subjects tended to be much more willing to sacrifice the fat man, even though this still meant using the fat man to save the others.

The embodied factor that Greene and colleagues identified as relevant is called *personal force*.[31] The intuitions thought to support deontology, because it is elicited in a situation where a person is used as a mere means to an end, actually pick up on a fairly direct bodily interaction towards another person (experiments support that this is both the case for pushing with one's hands and using a pole). As it turns out, what a subject's decisions depend on is whether they have to execute specific movement patterns. Pulling a lever is unproblematic for subjects, while shoving or pushing motions are problematic.[32] The issue that these experiments raise is that this seems to be a morally irrelevant factor, not in the sense that *personal force* is morally acceptable, but that this is not a suitable criterion for differentiating between morally acceptable and unacceptable behavior; if someone takes action to harm another human, it should not depend on the manner in which this harm is inflicted whether this is acceptable or not.

The issue for moral sentimentalists arises from the putative connection between emotional reactions and these embodied *personal force* judgements. Given the cases which Greene investigates such a link is reasonable to conclude. While, he appears to absorb embodied aspects of cognition into the fast-automatic portion of his dual-process moral brain, usually embodiment theorists would not restrict their claims about the embodied nature of the mind to affective phenomena. However, the connection between emotions and embodiment cognitive processes itself is

plausible given a number of recent works asserting similar claims.[33] This connection is problematic, because an embodied (or maybe even more generally situated) system underlying moral judgements may have some troubling implications. Wilson and Foglia define embodied cognition as follows:

> *Embodiment thesis:* Many features of cognition are embodied in that they are deeply dependent upon characteristics of the physical body of an agent, such that the agent's beyond-the-brain body plays a significant causal role, or a physically constitutive role, in that agent's cognitive processing.[34]

In light of Greene and colleagues' insights this can be adapted to an *Embodied Moral Cognition Thesis* (EMCT):

> Many features of moral cognition (especially emotional ones) are dependent upon characteristics of the physical body of an agent.

EMCT implies that emotional moral responses are not only relative to one agent's beliefs about the world, but to the bodily changes they experience. This specific form of subjectivism – grounded in an agent's body – may for some sentimentalist be the proverbial straw that breaks the camel's back. An inconsistency argument is easily constructed:

- Theory S holds emotions to be morally relevant, because they rely on non-embodied factors X, Y, or, Z (e.g., on them being related to beliefs, language, etc.), which make emotions sensitive to morally relevant states of affair.
- Emotional process E is causally responsible for responses to certain moral problems.
- Experimental evidence suggests that process E actually relies on embodied motor patterns, that are not a suitable cognitive vehicle for factors X, Y, and Z.
- Thus, theory S defending the moral significance of emotions on the grounds that they rely on factors X, Y, or Z has been debunked.

Such a debunking argument is only applicable to a theory of (moral) emotions stipulating non-embodied factors as constitutive for (moral) emotions. EMCT strongly suggests certain meta-ethical conclusions, for example, that emotional moral thought does not resemble language, refer to beliefs, etc., but moral thought relies on (learned) sensory-motor patterns. Whether this has meaningful implications depends on the concrete concepts of cognitive processes in question. Traditionally, mental mechanics are conceived of as relying on their representational content about the world. But if the mental processes in question are sensory-motor patterns and presumably lack propositional representational content; traditional conceptions of moral thought, belief, and language, as well as moral theories relying on such conceptions, are called into question.[35]

It is common to conceptualize moral terms by the standards of universality, impartiality or objectivity. After all, they are supposed to be not only meaningful to the agents experiencing or expressing them, they should apply to all moral agents. Cognitivists tend to account for this by conceptualizing cognitive processes as being representational of true moral concepts existing in the world, referenced by all moral agents. The loss of representational content as a conceptual tool to account for these desiderata leaves a gap in many theories that needs to be closed. This will either require some sort of moral bridging principle that connects the bodily specific sensory-motor patterns to impartial, universalizable moral principle, or the abandonment of such principles.

To abandon such principles may not be farfetched, as a growing number of ethicists appear to move away from these principles and towards a more relational approach to ethics,[36] which may match up well with these cognitive insights.

This embodiment issue is not new to moral sentimentalists and emotion theorists, but still a contentious one. Embodied emotion theses have been around for a long time, e.g., the James-Lange-Theory, which construe emotions to be feelings of bodily change. Embodied (or situated) theories of emotions are historically the default positions.[37] But James-Lange inspired theories have been criticized for failing to account for the intentional content of emotions.[38] Meaningful moral emotions especially rely on their intentional content. Non-intentionality would invalidate moral sentimentalism, as emotions would motivate behavior in the world without meaningfully being about what actually happens in the world.

In the history of the philosophical study of emotions, this led to the advent of cognitivist theories of emotion. Theorists distanced themselves, or have been understood as distancing themselves, from the embodied aspects of emotions and reconceptualized emotions as cognitive processes, which carry intentional content in virtue of their representational content of the world. Authors conceived of emotions as judgements,[39] e.g., that fear is the judgement that something is dangerous. This tendency to reconceptualize emotions extended to the study of moral emotions, where moral-cognitivists conceptualize emotions as emotional judgements that contain, refer to or resemble moral beliefs or moral statements. These positions are prima facie inconsistent with embodied conceptions of (moral) emotions, as emotional responses (which would be labeled as judgements in these views) are based on sensory motor-patterns and thus would not contain, refer to or resemble moral beliefs or moral statements. Such views would need to be defended against such an EMCT argument.

But cognitivist conceptions of emotion have already fallen somewhat out of favor, based on the general evidence from cognitive science, that cognitive processes and emotions actually are in many ways embodied. And as Prinz[40] points out, the embodiment of emotions is not actually inconsistent with them being intentional. Increasingly, current theorists account for them as such, resolving the tension between intentionality and embodiment.[41]

Moral sentimentalist positions that are untroubled by the EMCT thesis are also easy enough to find. Most prominently Jesse Prinz advanced a theory of emotion and morality.[42] This theory strongly emphasizes embodied aspects of emotions and the moral significance of emotions. Furthermore, Colombetti and Torrance[43] provide a more substantial sketch of a primarily situated, i.e., enactivist, theory of morality, which is compatible with moral sentimentalism. Urban[44] advances this line of thought to connect enactivist ethics with care ethics, a modern variant of relational ethics. The connection to care ethics is especially appealing because as opposed to the comparatively ancient theories usually invoked in moral cognition debates, care-ethics emerged from early psychological enquiries into morality.[45]

Conclusion

The question investigated in the paper is whether moral cognition research is a threat to the position of moral sentimentalism. While this appears to be a somewhat common worry (or hope), it is not the case. Debunking arguments that aim to invalidate moral sentimentalism fall into two categories of arguments: sourcing arguments, i.e., debunking intuitions by reference to the neuropsychological source of the argument, or inconsistency arguments, i.e., invalidating intuitions by showing an experimentally revealed inconsistency between such intuitions and the empirical commitments of a theory.

In the first part of this paper, sourcing arguments against moral sentimentalism were revealed to be circular and thus not a valid form of argument. Sourcing arguments employ an implicit normative first principle to establish the neuropsychological source of an intuition as "good" or "bad." However, this principle, in turn, becomes the object of experimental investigation. The principle is con-

firmed (with some obfuscation) by showing that intuitions supporting this principle have a "good" source and intuitions supporting other principles have a "bad" source. This, however, is circular, because the principle established which source is "good" and "bad" in the first place, and thus the truth of the tacitly assumed principle, depends on the conclusion of the argument. To put it bluntly, this argument claims that moral sentimentalism is wrong because emotions rely on unreliable cognitive processes, but the unreliability has been established on rationalistic grounds, i.e., considering an emotional response to be unreliable in the first place. Without the argumentative force of sourcing arguments, moral cognition research appears toothless in its attacks on moral sentimentalism, though not irrelevant for moral sentimentalist.

The other species of debunking arguments, i.e., inconsistency arguments, are an interesting addition to the debate. Moral sentimentalists have to make certain empirical stipulations about the nature of emotions, which supply some form of cognitive contribution to moral cognizing but still mainly rely on the tools of philosophical intuitionism to defend their theories. It is an argument worth making as it shows that these concrete empirical stipulations are inconsistent with the experimentally investigated intuitions ostensibly supporting the same position. This has often been cashed out in the form of arguments against deontology, which somewhat misses the mark. While the presence of emotional factors in moral cognition is a problem for (some) deontologists, it is certainly not a problem for sentimentalists. Moral sentimentalism is supported by experimental evidence that shows emotions play a crucial role in moral cognizing; after all, sentimentalists usually assume this to be the case anyway.

However, the specifics of experimental evidence accumulated may not make all sentimentalists equally happy. As Greene and colleagues show,[46] there is persuasive evidence that moral cognition (and likely emotional processes especially) rely on a form of embodied cognition. The involvement of the body in moral emotions has been objected to, in the philosophical debate about the nature of emotions, on the grounds that this may preclude the intentionality of emotions. However, many modern conceptions of emotions treat the embodiment and intentionality of emotions as compatible. Accordingly, there are sentimentalist positions that are consistent with experimental evidence suggesting the embodiment of emotions. Consequently, research from moral cognition is not a threat to moral sentimentalism, but a welcome empirical source of information that enriches and substantiates the debate.

Acknowledgements

This publication is funded by the DFG-Graduiertenkolleg "Situated Cognition", GRK-2185/1.

Notes

[1] Cf. E.E. WILSON, L. DENIS, *Kant and Hume on Morality*, in E.N. ZALTA (ed), *The Stanford Encyclopedia of Philosophy*, Summer Edition 2018, URL: <https://plato,stanford,edu/archives/sum2018/entries/kant-hume-morality/>.

[2] A. KAUPPINEN, *Moral Sentimentalism*, in: E.N. ZALTA (ed), *The Stanford Encyclopedia of Philosophy*, Spring Edition 2017, URL: <https://plato,stanford,edu/archives/spr2017/entries/moral-sentimentalism/>.

[3] Cf. J.D. GREENE, R.B. SOMMERVILLE, L.E. NYSTROM, J.M. DARLEY, J.D. COHEN, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, in: «Science», vol. CCXCIII, n. 5537, 2001, pp. 2105-2108; J.D. GREENE, F.A. CUSHMAN, L.E. STEWART, K. LOWENBERG, L.E. NYSTROM, J.D. COHEN, *Pushing Moral Buttons: The Interaction between Personal Force and Intention in Moral Judgment*, in: «Cognition», vol. CXI, n. 3, 2009, pp. 364-371; J.D. GREENE, *Moral Tribes: Emotion, Reason, and the Gap between Us and Them*, Penguin, New York 2014; J.D. GREENE, *Beyond Point-and-Shoot Morality: Why Cognitive (Neuro) Science Matters for Ethics*, in: «Ethics», vol. CXXIV, n. 4, 2014, 695-726. For a similar argument see P. SINGER, *Ethics and Intuitions*, in: «The Journal of Ethics», vol. IX, n. 3-4,

2005, pp. 331-352.

[4] Cf. R. RAILTON, *The Affective Dog and its Rational Tale: Intuition and Attunement*, in: «Ethics», vol. CXXIV, n. 4, 2014, pp. 813-859; C. KLEIN, *The Dual Track Theory of Moral Decision-making: A Critique of the Neuroimaging Evidence*, in: «Neuroethics», vol. IV, n. 2, 2011, pp. 143-162; S. BERKER, *The Normative Insignificance of Neuroscience*, in: «Philosophy & Public Affairs», vol. XXXVII, n. 4, 2009, pp. 293-329; J. MOLL, R. DE OLIVEIRA-SOUZA, R. ZAHN, *The Neural Basis of Moral Cognition: Sentiments, Concepts, and Values*, in: «Annals of the New York Academy of Sciences», vol. MCXXIV, 2008, pp. 161-180; R. SUN, *Moral Judgement, Human Motivation, and Neural Networks*, in: «Cognitive Computation», vol. V, n. 4, 2013 , pp. 566-579; G. KAHANE, J.A. EVERETT, B.D. EARP, L. CAVIOLA, N.S. FABER, M.J. CROCKETT, J. SAVULESCU, *Beyond Sacrificial Harm: A Two-dimensional Model of Utilitarian Psychology*, in: «Psychological Review», vol. CXXV, n. 2, 2018, pp. 131-164; P. KÖNIGS, *On the Normative Insignificance of Neuroscience and Dual-process Theory*, in: «Neuroethics», vol. XI, n. 2, 2018, pp. 195-209.

[5] Cf. J.D. GREENE, J. HAIDT, *How (and Where) does Moral Judgment Work?*, in: «Trends in Cognitive Sciences», vol. VI, n. 12, 2002, pp. 517-523; J. HAIDT, S.H. KOLLER, M.G. DIAS, *Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog*, «Journal of Personality and Social Psychology », vol. LXV, n. 4, 1993, pp. 613-628.

[6] Whether the responses to experimental stimuli have the same force as philosophical intuitions is questionable. However, as this paper is concerned with the validity of the arguments and not with the correctness of the premises, it will charitably accept this premise.

[7] My use of the term "intuition" differs from Greene's. Greene supposes intuitions to be automatic emotional responses. These emotional intuitions form one half of his dual process morality, supplying a common sense morality. This common sense morality is more or less what people rely on in everyday social life, and it works reasonably well in these situations. I use "intuition" in a broader sense: They are not a species of either automatic or controlled processing (Wan and colleagues provide neurophysiological evidence that intuition are also a feature of controlled processing; see X. WAN, D. TAKANO, T. ASAMIZUYA, C. SUZUKI, K. UENO, K. CHENG, T. ITO, K. TANAKA, *Developing Intuition: Neural Correlates of Cognitive-skill Learning in Caudate Nucleus*, in: «Journal of Neuroscience», vol. XXXIII, n. 48, 2012, pp. 17492-17501). I use "intuition" more closely to the philosophical meaning of an intuition rather than the psychological notion Greene is working with. Thus my use of this term is supposed to be without tacit normative assertions (as opposed to Greene supposing intuitions to be an inferior source of moral insight).

[8] Cf. A. KAUPPINEN, *Moral Sentimentalism*, cit.

[9] Cf. P. SINGER, *Ethics and Intuitions*, cit.; J.D. GREENE, F.A. CUSHMAN, L.E. STEWART, K. LOWENBERG, L.E. NYSTROM, J.D. COHEN, *Pushing Moral Buttons*, cit.; J.D. GREENE, S.A. MORELLI, K. LOWENBERG, L.E. NYSTROM, J.D. COHEN, *Cognitive Load Selectively Interferes with Utilitarian Moral Judgment*, in: «Cognition», vol. CVII, n. 3, 2008, pp. 1144-1154; J.D. GREENE, *Moral Tribes*, cit.; J.D. GREENE, *Beyond Point-and-Shoot Morality*, cit.

[10] J.D. GREENE, *Beyond Point-and-Shoot Morality*, cit., p. 715.

[11] Cf. J.D. GREENE, F.A. CUSHMAN, L.E. STEWART, K. LOWENBERG, L.E. NYSTROM, J.D. COHEN, *Pushing Moral Buttons*, cit.; J.D. GREENE, *Moral Tribes*, cit.; J.D. GREENE, *Beyond Point-and-Shoot Morality*, cit.; J.D. GREENE, *The Rat-a-gorical Imperative: Moral Intuition and the Limits of Affective Learning*, in: «Cognition», vol. CLXVII, 2017, pp. 66-77.

[12] Premise (I) can also be formulated differently to provide the same argumentative force to the argument, e.g., Greene provides an alternative defense of this point drawing models of the neurological learning mechanisms employed by these processes (see J.D. GREENE, *The Rat-a-gorical Imperative*, cit.). I contend that this more elaborate defense only complicates the issue and will not provide a more defensible position. The main point is that different intuitions are sourced from neurological processes with differing functional profiles. The issue raised in the next section, that normativity is assigned to these different processes in an invalid way to construct a sourcing argument, holds irrespective of how specifically these processes are described.

[13] J.D. GREENE, *Moral Tribes*, cit., pp. 119-131.

[14] Cf. J.D. GREENE, R.B. SOMMERVILLE, L.E. NYSTROM, J.M. DARLEY, J.D. COHEN, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, cit.

[15] Cf. P. FOOT, *The Problem of Abortion and the Doctrine of Double Effect*, in: «Oxford Review»,

vol. V, 1967, pp. 5-15.

[16] Cf. J.J. THOMSON, *Killing, Letting Die, and the Trolley Problem*, in: «The Monist», vol. LIX, n. 2, 1976, pp. 204-217; J.J. THOMSON, *The Trolley Problem*, in: «The Yale Law Journal», vol. XCIV, n. 6, 1985, 1395-1415.

[17] Kahane and colleagues rightly raise the issue that this may not be the case, as subjects may still embrace conflicting views on other aspects of a philosophical position (see G. KAHANE, J.A. EVERETT, B.D. EARP, L. CAVIOLA, N.S. FABER, M.J. CROCKETT, J. SAVULESCU, *Beyond Sacrificial Harm*, cit.).

[18] Cf. J.D. GREENE, R.B. SOMMERVILLE, L.E. NYSTROM, J.M. DARLEY, J.D. COHEN, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, cit.

[19] Cf. D. KAHNEMAN, *Thinking, Fast and Slow*, Farrar, Straus & Giroux, New York 2011.

[20] Cf. J.S.B. EVANS, *Hypothetical Thinking: Dual Processes in Reasoning and Judgment*, Psychology Press, 2007; J.S.B. EVANS, *Dual Process Theories of Deductive Reasoning: Facts and Fallacies*, in: K.J. HOLYOAK, R.G. MORRISON (eds.), *The Oxford Handbook of Thinking and Reasoning*, Oxford Scholarship, Oxford 2012, pp. 115-133; J.S.B. EVANS, *Questions and Challenges for the New Psychology of Reasoning*, in: «Thinking & Reasoning», vol. XVIII, n. 1, 2012, pp. 5-31.

[21] J.S.B. EVANS, *Dual Process Theories of Deductive Reasoning*, cit., p. 127 - Square brackets indicate clarifications by the author.

[22] *Ibidem*.

[23] *Ibidem*.

[24] *Ibidem*.

[25] D. KAHNEMAN, *Thinking, Fast and Slow*, cit., pp. 39-49.

[26] *Ivi*, pp. 156-165.

[27] I take this to be a somewhat reasonable approximation, of the implicit rationalistic assumptions diagnosed in section 4. A mathematical feature of the world would for example be that 5 is more than 1 or that a ball a dollar cheaper than a bat costing 1.05$ costs 0.05$. Such mathematical features may have moral relevance if moral considerations are made in term of minimizing deaths, e.g., in the trolley problem that there are less people on one track than the other.

[28] Cf. E. MIHAILOV, *Is Deontology a Moral Confabulation?*, in: «Neuroethics», vol. IX, n. 1, 2016, pp. 1-13; A. KAUPPINEN, *Moral Sentimentalism*, cit.

[29] To assert the moral irrelevance of these factors may be prima facie plausible. But it should be considered, that a moral theory committed to a claim like "agents should not pick up on factor X" is committed to the claim "agents are able to not pick up on factor X" (should implies can). The empirical evidence, however, suggests that it is a distinct possibility that emotions or embodied processes (see the following sections) are integral parts of moral cognition and cannot just be not picked up on. Thus, theories committed to such claims are in some trouble.

[30] J.D. GREENE, F.A. CUSHMAN, L.E. STEWART, K. LOWENBERG, L.E. NYSTROM, J.D. COHEN, *Pushing Moral Buttons*, cit., p. 370.

[31] *Ibidem*.

[32] *Ibidem*. See also J.D. GREENE, *Moral Tribes*, cit., p.124 and p. 215.

[33] Cf. P. GRIFFITHS, A. SCARANTINO, *Emotions in the Wild: the Situated Perspective on Emotion*, in: P. ROBBINS, M. AYDEDE (eds,), *The Cambridge Handbook of Situated Cognition*, Cambridge University Press, Cambridge 2009, pp. 437-453; J.J. PRINZ, *Gut Reactions: A Perceptual Theory of Emotion*, Oxford University Press, Oxford / New York 2004; M, MAIESE, *Embodiment, Emotion and Cognition*, Palgrave Macmillan, Basingstoke 2011; G. COLOMBETTI, *The Feeling Body. Affective Science meets the Enactive Mind*, MIT Press, Cambridge (MA) 2014; A. STEPHAN, S. WALTER, W. WILUTZKY, *Emotions Beyond Brain and Body*, in: «Philosophical Psychology», vol. XXVII, n. 1, 2014, pp. 65-81; R. HUFENDIEK, *Embodied Emotions: A Naturalist Approach to a Normative Phenomenon*, Routledge, London / New York 2015.

[34] R.A. WILSON, L. FOGLIA, *Embodied Cognition*, in: E.N. ZALTA (ed.), *The Stanford Encyclopedia of Philosophy* (Spring edition 2017), URL: <https://plato,stanford,edu/archives/spr2017/entries/embodied-cognition/>.

[35] Of course there are embodied conceptions of these cognitive processes (see, e.g., D. CASASANTO, *Embodiment of Abstract Concepts: Good and Bad in Right- and Left-handers*, in: «Journal of Experimental Psychology: General», vol. CXXXVIII, n. 3, 2009, pp. 351-367; D. CASASANTO, *Different Bodies, Different Minds: The Body Specificity of Language and Thought*, in: «Current Directions in Psychological Science», vol. XX, n. 6, 2011, pp. 378-383), but whether they retain the elements moral theories rely on is a matter for a more comprehensive investigation.

[36] Cf. T. METZ, S.C. MILLER, *Relational Ethics*, in:

The International Encyclopedia of Ethics, Wiley Online Library, Malden (MA) 2016, p. 6 - doi: 10.1002/9781444367072.wbiee826.

[37] Cf. W. WILUTZKY, A. STEPHAN, S. WALTER, *Situierte Affektivität*, in: J. SLABY, A. STEPHAN, H. WALTER, S. WALTER (Hrsg.) *Affektive Intentionalität: Beiträge zur welterschliessfunktion der Emotionen*, Mentis, Paderborn 2011, pp. 283-320.

[38] Cf. J.J. PRINZ, *Emotions Embodied*, in: R. SOLOMON (ed.), *Thinking about Feeling*, Oxford University Press, Oxford / New York 2004, pp. 44-60.

[39] Cf. M.C. NUSSBAUM, *Upheavals of Thought. The Intelligence of Emotions*, Cambridge University Press, Cambridge 2001; R. SOLOMON, *Emotion and Choice*, in: A. RORTY (ed.), *Explaining Emotions*, University of California Press, Los Angeles (CA) 1980, pp. 251-281.

[40] Cf. J.J. PRINZ, *Emotions Embodied*, cit.

[41] Cf., e.g., J. SLABY, A. STEPHAN, *Affective Intentionality and Self-consciousness*, in: «Consciousness and Cognition», vol. XVII, n. 2, 2008, pp. 506-513; P. GRIFFITHS, A. SCARANTINO, *Emotions in the Wild: the Situated Perspective on Emotion*, cit.; G. COLOMBETTI, *The Feeling Body*, cit.; R. HUFENDIEK, *Embodied Emotions*, cit.

[42] Cf. J.J. PRINZ, *The Emotional Construction of Morals*, Oxford University Press, Oxford / New York 2007.

[43] Cf. G. COLOMBETTI, S. TORRANCE, *Emotion and Ethics: An Inter-(en)active Approach*, in: «Phenomenology and the Cognitive Sciences», vol. VIII, n. 4, 2009, pp. p. 505.

[44] Cf. P. URBAN, *Enactivism and Care Ethics: Merging Perspective*, in: «Filozofia», vol. LXX, n. 2, 2015, pp. 119-129; P. URBAN, *Enacting Care*, in: «Ethics and Social Welfare», vol. IX, n. 2, 2015, pp. 216-222; P. URBAN, *Foregrounding the Relational Domain. Phenomenology, Enactivism and Care Ethics*, in: «Horizon. Studies in Phenomenology», vol. V, n. 1, 2016, pp. 171-182.

[45] Cf. J. PIAGET, *The Moral Judgment of the Child* (1932), Free Press, New York 1965; L. KOHLBERG, Moral Stages and Moralization: The Cognitive-development Approach, in: «Infancia y Aprendizaje», vol. V, n. 18, 1976, pp. 31-53; C. GILLIGAN, *In a Different Voice*, Harvard University Press, Cambridge (MA) 1982.

[46] Cf. J.D. GREENE, F.A. CUSHMAN, L.E. STEWART, K. LOWENBERG, L.E. NYSTROM, J.D. COHEN, *Pushing Moral Buttons*, cit.